The Interplay between Feature-Saliency and Feedback Information in Visual Category Learning Tasks

Rubi Hammer[1] (rubih@stanford.edu)
Vladimir Sloutsky[2] (sloutsky@psy.ohio-state.edu)
Kalanit Grill-Spector[1] (kalanit@stanford.edu)
1. Department of Psychology, Stanford University
2. Department of Psychology and Center for Cognitive Science, Ohio-State University

Abstract

What is the role of feedback information in different visual category learning (VCL) scenarios? To address this question we tested participants' performance in VCL tasks in which stimuli varied in three feature dimensions, one of which was relevant for the task and the other two were irrelevant. The relevant feature could be identified based on trial-by-trial feedback. In one condition the task relevant and irrelevant features were highly-salient. In the second condition all features had low-visual-saliency. Feedback information was also manipulated: In the high-information condition the task relevant feature could be identified by the information provided in each trial whereas in the mid-information condition the feedback was ambiguous and information from several learning trials was required in order to confidently identify the relevant feature. Surprisingly, our data shows that mid- and high-information feedback are similarly effective in high-saliency VCL tasks. In contrast, in low-saliency VCL tasks, mid-information feedback impairs learning. We suggest that VCL can be done effectively either when feedback is ambiguous or in low-saliency conditions, but not in scenarios when both challenges occur concurrently.

Keywords: Visual category learning; Feedback information; Attentional learning; Perceptual learning; Feature-saliency.

Introduction

Humans are capable of effectively managing a vast amount of sensory information, rapidly rendering it into a coherent, reliable and meaningful representation of objects and events. This capability depends on two fundamental learning processes: One is perceptual learning which allows identifying subtle, initially hard to detect, differences between stimuli (Goldstone & Barsalou, 1998; Kourtzi, 2010). The other is attentional learning which requires shifting attention to relevant attributes while at the same time filtering out irrelevant, even if salient, visual attributes (Blair, Watson & Meier, 2009; Kalish & Kruschke, 2000; Rehder & Hoffman, 2005). It is known that these two forms of learning allow reducing the probability of future decision errors by improving discriminability among similar objects from different categories, and allowing effective generalization to novel stimuli. To date, it is not clear how these processes interact in different learning scenarios.

Difficulty in perceptual learning tasks is determined by feature-saliency. Difficulty in attentional learning tasks is determined by the numbers of simultaneous perceptual features one has to process when categorizing objects.

Learning trajectories in both attentional and perceptual learning tasks are also determined by the availability of informative feedback. In this study we examine the interaction between perceptual and attentional learning by testing the interaction between feature-saliency and feedback information in visual category learning (VCL) tasks of complex visual stimuli.

We define feature-saliency as the physical dissimilarity between stimuli in a given feature dimension. When objects are perceived as highly dissimilar across a feature dimension, this feature is perceived as more diagnostic than lower-saliency feature dimensions (Chin-Parker & Ross, 2004). In each VCL task in our experiment stimuli differed from one another in three feature dimensions, yet only one was relevant for correctly categorizing the stimuli. In each task we kept the relative feature saliency (low or high) similar across all three feature dimensions making them equal candidates for being perceived as task relevant. Therefore, the diagnostic value of each feature could be determined only by the information provided by feedback.

We define the feedback information level based on its ambiguity. In each learning trial we presented a pair of stimuli and the participant had to determine if these stimuli belong to the same category or different categories. We used two levels of feedback information: 1) In the *high-information* learning condition the feedback in each and every trial provided sufficient information for learning the rule as *same-category pairs* were identical only in the task relevant feature and differed in the two irrelevant ones, whereas *different-categories pairs* were different only in the task relevant feature (and identical in the two irrelevant ones). If A denotes the relevant feature, B and C the irrelevant features, and X is the outcome of a categorization decision, the only possible trial-by-trial inferred causality (a feasible feature-decision association) in the high-information condition was $A \rightarrow X \cap A \rightarrow X \cap A \rightarrow X \dots$ 2) In the *mid-information* condition each trial was ambiguous as *same-category pairs* were identical in the task relevant feature and one of the two irrelevant features (randomly alternating between the two), whereas d*ifferent-categories pair* were different in the task relevant feature and one of the irrelevant features. The causality here was $(A \cup B \rightarrow X) \cap (A \cup C \rightarrow X) \cap (A \cup C \rightarrow X) \cap (A \cup B \rightarrow X) \dots$ This learning scenario is more likely to require distributing attention between features and integrating information across more trials in order to learn the categorization rule.

We expect VCL efficiency to depend both on feature-saliency and the level of feedback information. Specifically, we hypothesize that when differences across visual feature dimensions are hard to detect, shifting attention from one feature to the other will not be effective due to the poor representation of features. This will become a greater challenge when feedback is ambiguous, making it harder to associate a feature representation with the corresponding decision outcome (Nosofsky & Palmeri, 1996). In contrast, when feature-saliency is high, relevant visual information is readily available, enabling to rapidly shift attention away from irrelevant features. This may allow effective learning even in the face of ambiguous feedback. Therefore, we expect an interactive effect of feature-saliency and feedback information. Nevertheless, based on current knowledge, we cannot predict whether this interaction will manifest as an additive effect or as an augmented interference in which the effect of feedback ambiguity on learning efficiency will be more profound in low-saliency learning conditions.

Methods

Participants

Sixty paid adults (36 females), with normal or corrected to normal vision, participated in the experiment. The experiment was approved by the Stanford University IRB.

Materials and setting

We ran the experiment using Psychtoolbox (MATLAB®) on 1920X1200 pixels computer display. Participants' head was located about 70 cm (~2 feet) from the computer screen such that each one of the two simultaneously presented stimuli occupied approximately 14° of the visual field.

Stimuli

We used four distinct sets of novel creature-like stimuli. In each set the stimuli varied in three feature dimensions (see examples in Figure 1). Exemplars for each stimulus set were produced from one standard object and three morph targets, each differed from the standard in one feature dimension. For VCL tasks with high-feature-saliency we used high morph values (at least 77%). Low-saliency exemplar pairs differed by small morph values (22-33%). The morphing values we used were determined based on pilot tests such that within each stimulus set differences in all feature dimensions were similarly likely to be detected. We also ensured that in the low-saliency condition differences within each feature dimension will be not detected easily without feedback. For each stimulus set we determined two categories. Members of each category varied in the two irrelevant feature dimensions and were identical in the third. This third feature-dimension was the diagnostic feature-dimension differentiating between the two categories.

Design

Tasks differed in feature saliency (high-saliency vs. low-saliency), and three levels of feedback information. In high-saliency VCL tasks both within-category and between-categories differences had high-saliency. In low-saliency VCL tasks both within-category and between-categories differences had low-saliency. In each VCL task participants were trained with one of three levels of feedback information: 1) High-level information feedback that potentially enabled identifying the diagnostic feature within each learning trial; 2) mid-level information feedback, in which each learning trial provided ambiguous information regarding which feature was the relevant one; 3) In addition to these two feedback-based learning conditions, we also tested participants in a control condition in which no feedback was provided to the participants. This provided a useful benchmark for assessing the contribution of feedback information to learning.

Each VCL task was based on a different stimulus set and a unique combination of feature-saliency and feedback information level. In order to prevent cross-conditions differences in categorization performances that derive from differences between stimulus sets, we counterbalanced the tasks across participants such that each one of the four stimulus sets was used in each one of the six experimental conditions the same number of times.
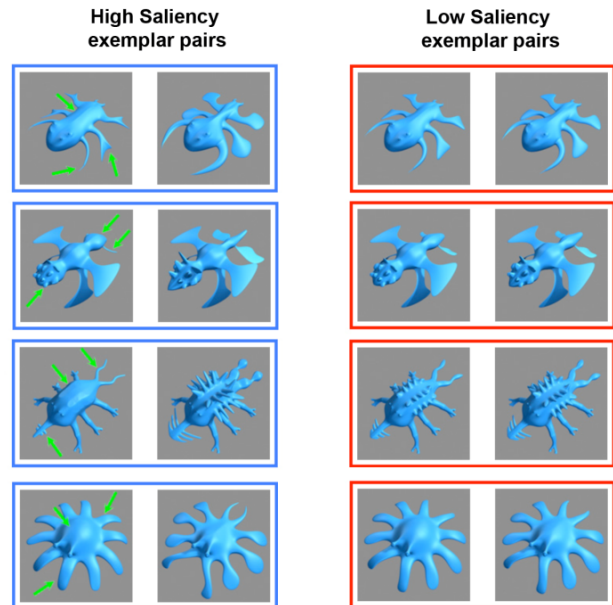

Figure 1: Examples of pairs of creatures from the four stimulus sets. Each row shows examples from a different stimulus set. *Left:* high-saliency pairs; *Right:* low-saliency pairs. Green arrows indicate high-saliency differences in the three feature dimensions in which the paired creatures differ (e.g., in the upper row the two creatures differ in horns, limbs and body-with). In each stimulus set, the same diagnostic feature was used for both the low- and high-saliency tasks. From top to bottom, the task relevant feature was body-width, head-spikes, horns and body curvature.

Each VCL task included seven blocks: four test blocks (denoted as T1-T4) that alternated with three learning blocks (L1-L3). Each block consisted of 24 trials. In each

trial two creatures were presented simultaneously on a computer screen for 2.2 seconds during which the participant had to decide if they belonged to the same or different categories by pressing one of two keyboard keys. Feedback was given during the 0.8 seconds inter-trial interval: In the mid- and high-information feedback conditions, a green square indicated a correct answer and red square an incorrect one. In the control, no feedback, learning blocks and in the test blocks, a yellow square indicated that the response was recorded (Figure 2a).

During test blocks, creature pairs always differed in two feature dimensions. *Same-category pairs* differed in the two irrelevant feature dimensions and were identical in the relevant one (as the right upper pair in Figure 2b). *Different categories pairs* differed in the relevant feature dimension and in one of the irrelevant feature dimensions, but were identical in the other irrelevant dimension (see left lower pair in Figure 2b). This design prevents participants from making the same/different categorization decision based on the overall similarity between stimuli. It also allowed us to keep the statistics of the three features and their pair-wise covariance identical.

For feedback-based learning blocks, pairs of creatures were selected in the following way: 1) In the high-information feedback condition *same-category pairs* were identical in the relevant feature dimension and differed in the two irrelevant feature dimensions. *Different-categories pairs* differed in the relevant feature dimension and were identical in the two irrelevant ones. Thus, each trial indicates either all the within-category variability (*same-category pairs*), or only the diagnostic feature dimension discriminating between categories (*different categories pairs*). 2) For mid-information feedback, *same-category pairs* were identical in the relevant dimension and in one of the two irrelevant dimensions (alternating between the two across different trials). *Different-category pairs* differed in two features: the relevant one and an irrelevant one (again, randomly alternating across trials between the irrelevant two). Although in this case each trial was ambiguous, it was still possible to learn the categorization rule based on the information provided across several trials. Such ambiguity in the feedback keeps the attentional learning aspect of the task more challenging by increasing the probability that participants will divide attention, in each trial, between two feature dimensions that are equally perceived as relevant. 3) The composition of the trials in the learning blocks with no feedback was similar to the one used in the test blocks.

Procedure
To keep the duration of the experimental session short (~ 75 minutes), each participant performed a combination of three out of six conditions; [2 feature-saliency] X [3 feedback information]. Participants were informed that in each VCL task they have to learn to classify unfamiliar creatures from two distinct subspecies based on one attribute (feature dimension). Participants were also told that any variability in other attributes should be considered as irrelevant and ignored. Before starting the experimental tasks, participants performed a warm-up VCL task (with a different stimulus set). This enabled the participant to become familiarized with the experimental tasks.
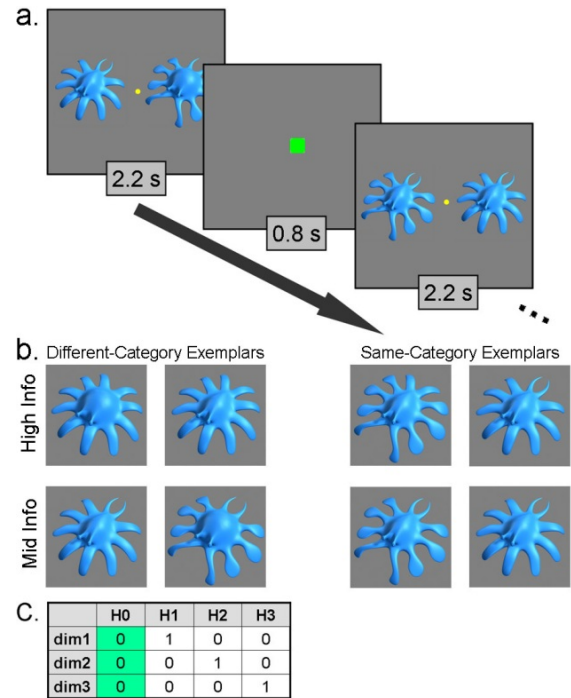

Figure 2: Experimental design. (a) An illustration of an experimental trial. A pair of stimuli presented for 2.2 seconds during which the participant had to judge if the creatures are from same or different categories. This was followed by 0.8 seconds of feedback presentation (e.g., green square indicates a correct answer) after which the next trial started. (b) Examples of different-categories (left) and same-category (right) pairs, high-information (top) and mid-information (bottom) condition. In each of the feedback based learning blocks, the category relation between the paired stimuli could be derived from the feedback. In High-information trials different-categories pairs differ only in the relevant feature, and same-category pairs are identical only in this feature. Such a trial enables effectively pinpointing the relevant feature dimension by eliminating 2 out of three possible hypotheses. Mid-informative trials provide less information since such a trial always leaves two options (out of three). (c) A table describing all possible hypotheses for a given VCL task. When all feature dimensions are salient the participant is only required to decide which feature is relevant (H1 – H3). When features are not salient, VCL requires shifting from H0 (represent a case in which the participant is unaware of either one of the potentially relevant feature dimensions) to the correct hypothesis signaling the relevant feature dimension.

Performance measurements
We define a "Hit" as correctly deciding that two creatures are members of the same category, and a "False-Alarm" as

incorrectly deciding that creatures of different categories are members of the same category. Based on the Hit and False-Alarm rate we calculated participants' sensitivity using the non-parametric measure A-prime (Grier, 1971; Formula 1). A'=1 indicates perfect performance and A'=0.5 indicates chance level. $0 \leq A' < 0.5$ represent a response confusion.

Formula 1: A-prime calculation. H denotes the Hit rate, and F the False-Alarm rate (Hit and False-Alarm rates are calculated based on the 24 trials in each test/learning block).

$$A' = 0.5 + \left[sign(H-F) \times \frac{(H-F)^2 + |H-F|}{4 \times \max(H,F) - 4 \times H \times F} \right]$$

Benchmarks
We evaluated participant performances according to the following benchmarks: Chance performance, A'=0.5; Perfect performance, A'=1; Performance based on systematically referring to an irrelevant feature during a test block or during a no-feedback "learning" block, A'= 0.12; Performance based on systematically referring to an irrelevant feature during a learning block with mid-information feedback, A'=0.5; Performance based on systematically referring to an irrelevant feature during a learning block with high-information feedback , A'= 0.

Results
Reported data is based on 24 participants in each condition. Excluded from this analysis are cases in which performance level was inconsistent between the test vs. the learning trials within a given VCL task (evident as high performance in the learning blocks, where feedback was available, contrasted with near-chance performance in the following test blocks, where feedback was not available).

Pre-learning performance
First, we confirmed that the initial performance level, in each of the two feature-saliency conditions, is similar. A two-way analysis of variance (ANOVA) with feature-saliency and feedback information as independent variables, and participants' sensitivity (A') in the first, pre-learning, test block (T-1) as the dependent variable shows no significant interaction $F_{(2, 144)} = 0.93$, no main effect of feedback information $F_{(2, 144)} = 0.91$, and no main effect of feature-saliency $F_{(1, 144)} = 2.12, p = 0.15$.

Trial-by-trial, feedback-based, learning dynamics
We assessed categorization improvement in the mid- and high-feedback information conditions based on learning trajectories across 72 learning trials (across the three learning blocks, L1 to L3). We calculated performance based on a moving average with a window of six trials. Figure 3 shows rapid learning, with almost identical trajectories, both with mid- and high-information feedback in the high-saliency condition. In contrast, in the low-saliency condition mid-information feedback resulted with significantly lower improvement compared with high-information feedback. Note that high-information feedback

ended with similarly high performance level in both saliency conditions, yet in the low-saliency condition it required more learning trials.

Separate two-way ANOVAs, one for the high-saliency condition and one for the low-saliency condition, with feedback information (mid/high) as the between participants independent variable, learning trial number (1-72) as a within participant independent variable, and participants' percent correct as the dependent variable, show a significant difference in the linear contrast between mid- and high-information feedback learning conditions in the low-saliency condition, $F_{(1, 46)} = 4.73, p < 0.04$, but not in the high-saliency condition, $F_{(1, 46)} = 0.29$.

Indeed, a three-way ANOVA with feature-saliency, feedback information and learning trial number as independent variables, and participants' percent correct as the dependent variable, confirm that the interaction between feature-saliency and feedback information is significant, $F_{(3, 92)} = 2.73, p < 0.02$, partial $\eta^2 = 0.029$ (Figure 3).
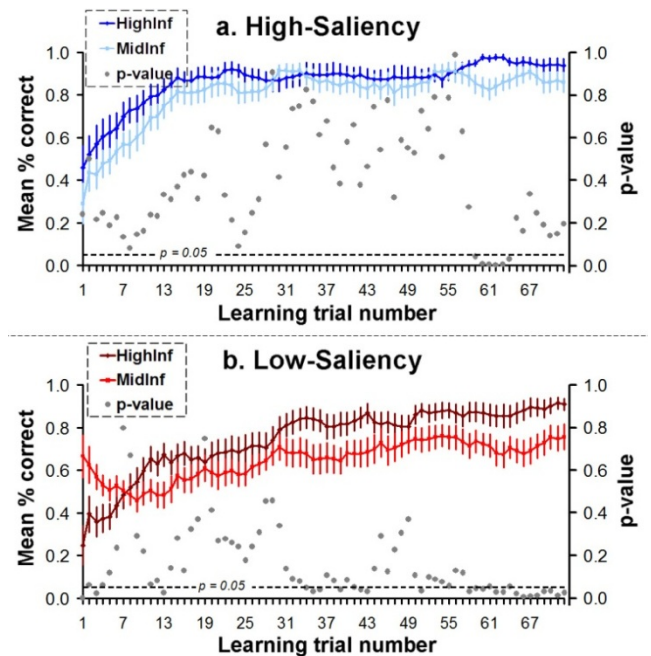

Figure 3: Participants' mean percent correct in the high-saliency (a) and low-saliency (b) conditions in each of the 72 learning trials. The value in each bin is based on moving average with a window of 6 trials (error bars represent standard error of the mean across 24 participants). Gray dots represent the significance level (p-value) of the difference between mid- and high-information learning in each trial (based on independent sample t-tests; Dashed line marks a significance level of $p = 0.05$). This illustrates a consistent significant difference between the mid- and high-information conditions, particularly in the second half of the learning process, only in the low feature-saliency condition.

Between test blocks dynamics
To further assess participants learning, we examined their performance in the test blocks where no feedback was given

and all trials had the same composition irrespective of the feedback information condition (paired creatures always differed in 2 features; see Methods). Results are shown in Figure 4. Our analysis shows significant improvement in all learning conditions (all $p < 0.01$). Importantly, we found a significant difference in the learning trajectories between the mid- and high-information feedback conditions in the low-saliency condition but not in the high-saliency condition.
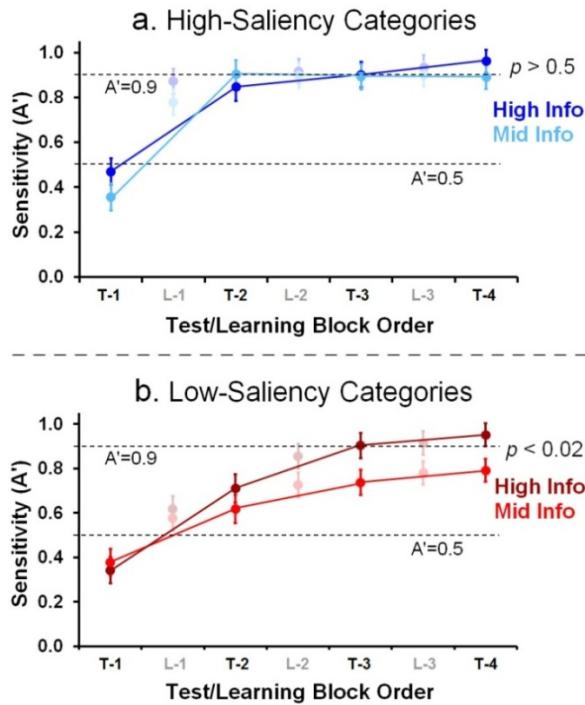

Figure 4: Participants' mean sensitivity (error bars represent standard error of the mean) in the (a) high-saliency and (b) low-saliency conditions. Feedback information levels are marked by different colors. T-1 to T-4: Performance in the test blocks (data points connected with lines); L-1 to L-3: Performance in the learning blocks. When feature-saliency is high, there are no significant differences in VCL efficiency between the mid- and high-information conditions. On the other hand, when the feature-saliency is low and participants were provided with mid-information feedback, performance was lower as compared with the high-information feedback condition.

High-feature-saliency VCL: A two-way ANOVA for the high-saliency condition, with feedback information and test block order (T-1 to T-4) as independent variables and participants' sensitivity (A') as a dependent variable shows no significant linear contrast between the mid- and high-information conditions, $F(1, 46) = 0.04$ (Figure 4a).

Low-feature-saliency VCL: In contrast to the above, in the low-saliency condition there was a significant linear contrast between the mid- and high-information feedback conditions, $F(1, 46) = 6.09$, $p < 0.02$ (Figure 4b).

Comparing high-feature-saliency and low-feature-saliency VCL: A three-way ANOVA with feature-saliency

(low/high), feedback information (mid/high) and test block order (T1 to T4) as independent variables, and participants' sensitivity (A') as a dependent variable, confirm that the interaction between feature-saliency and feedback information is close to significant, $F(3, 92) = 2.84$, $p = 0.06$, partial $\eta^2 = 0.030$.

Next, we examined if differences in sensitivity between the high- and low-saliency conditions are driven mostly by participants' Hit rate or by their False-Alarm rate. We found no significant effect in the False-Alarm rate and a trend in Hit rate: Two-way ANOVAs, one for the high-saliency and one for the low-saliency condition, with feedback information and test block order as independent variables and participants' Hit rate (in the test blocks) as the dependent variable, show a close to significant difference between the mid- and high-information feedback conditions in the low-feature-saliency condition, $F(1, 46) = 3.65$, $p = 0.06$, but not in the high-feature-saliency condition, $F(1, 46) = 0.57$. A similar analysis with False-Alarm rate as the dependent variable, shows no significant difference between mid- and high-information learning, neither in the low-saliency condition, $F(1, 46) = 1.49$, $p = 0.23$, or in the high-saliency condition, $F(1, 46) = 1.94$, $p = 0.17$.

These findings are surprising since the main challenge in low-saliency tasks is to learn to identify subtle important differences between similar categories (i.e. avoiding False Alarms) rather than deciding correctly that two apparently similar objects are from the same category (i.e. avoiding Misses). Nevertheless, our findings shows that higher information feedback in low-saliency conditions is mostly helpful in assisting participants performing better by avoiding Misses. We suggest that the lack of significant differences in False-Alarms rate represent, in fact, a response bias exhibited by participants in low saliency conditions – instead of discriminating between categories based on the relevant feature dimension, in the low-saliency tasks participants are more likely to react to any apparent subtle difference among paired creatures as if it is relevant, perhaps due to poor capacity in pinpointing the relevant one.

This interpretation is consistent with the apparent "superior" performance in the first few learning trials in the low-saliency mid-information condition (Figure 3b) where the participants seem to perform better than in the low-saliency high-information condition. In mid-information tasks, a strategy based on deciding "different categories" whenever identifying any difference, is with advantage since the task diagnostic feature is always coupled with an irrelevant one (whereas same-category pairs differ in only one, irrelevant, feature-dimension). That is, in low-saliency mid-information learning conditions people are likely to effectively avoid False-Alarms but for the wrong reason.

Performance in the control, no feedback, tasks
Finally, we confirmed that without feedback there is no significant learning in our VCL tasks: A two-way ANOVA conducted for the no feedback VCL tasks with feature-saliency (high/low) and test block order (T1 to T4) as

independent variable, and participants' sensitivity (A') as the dependent variable, shows no significant interaction between feature-saliency and test block $F(3, 46) = 1.73$, $p = 0.20$, and no test block order learning effect $F(3, 46) = 0.36$. Note that mean performance in the no-feedback tasks never significantly exceeded values of A'=0.5.

Discussion

We tested the interaction between feature-saliency and feedback information in visual category learning (VCL) tasks as a mean to explore the nature of the interaction between perceptual learning and attentional learning. Simply speaking, perceptual learning is a process that involves improvement in the ability to identify important fine differences between categories, whereas attentional learning improves the ability to filter out irrelevant (even if salient) within category differences (Hammer et al., 2009). Here we show that the interaction between these two processes is more complex than this simplistic view.

We report two important findings: First, perhaps surprisingly, we show that mid-information and high-information feedback are equally effective for learning when stimuli have marked visual differences as in the high-saliency condition. This suggests that when diagnostic visual information is readily accessible, ambiguity in feedback (which is associated with higher attention load) can be resolved with no apparent effort (at least in simple rule learning tasks and when testing typical adults). Second, importantly, there are substantial differences between mid-information and high-information feedback when stimuli are only subtly different. This suggests that low-saliency VCL depends more on informative feedback that serves to orient attention to the relevant feature. Unlike in high-information learning trials, in mid-information learning trials with low-saliency features, participants may not only face difficulties in noticing the relevant feature, but also have difficulties in disassociating it from irrelevant ones. Therefore, participants may have been unaware of the relationship between the feedback and the relevant feature-dimension, which consequently lowered learning effectiveness.

These findings are relevant to the developing debate on the role of attention in perceptual learning: Most findings suggest that perceptual learning requires attention to a target visual feature (Ahissar & Hochstein, 1993; Schoups et al., 2001), or the presence of informative feedback associated with an attended visual feature (Herzog & Fahle, 2002). In contrast, recent findings show that "accidental" perceptual learning can occur (Seitz, Kim & Watanabe, 2009). Nevertheless, this seems to be restricted to learning scenarios with informative feedback where unattended features are strongly correlated with an attended one.

Here we show that when there is only a partial positive correlation between the presentation of a task relevant feature and the presentation of irrelevant features (as it is inherently the case in mid-information feedback conditions), together with lack of explicit information regarding which feature is relevant, there is significant interference with the learning process. This is evident as significantly less effective learning compared with cases where the relevant feature and irrelevant features are consistently anticorrelated (as in high-feedback-information learning scenarios).

We conclude that the role of attention in visual learning tasks depends on the correlations between relevant and irrelevant features, the nature of information provided by available feedback, and feature-saliency. This suggests that in everyday life scenarios, when making judgments on complex objects in cluttered scenes, the relative contribution of attentional learning and perceptual learning can change quite substantially from one learning scenario to the other. Thus, perceptual learning and attentional learning should not be construed as mutually exclusive processes but rather as complementary processes, and visual learning tasks should be considered as a mixture of these two processes.

References

Ahissar, M. & Hochstein, S. 1993 Attentional control of early perceptual learning. PNAS, 90, 5718–5722.

Blair, M. R., Watson, M. R., & Meier. K. M. (2009). Errors, efficiency, and the interplay between attention and category learning. Cognition, 112(2), 330-336.

Chin-Parker, S., & Ross, B. H. (2004). Diagnosticity and prototypicality in category learning: A comparison of inference learning and classification learning. JEP: LMC, 30(1), 216-226.

Goldstone, R. L., & Barsalou, L. (1998). Reuniting perception and conception. Cognition, 65, 231-262.

Grier, J. B. (1971). Nonparametric indexes for sensitivity and bias: Computing formulas. Psychol Bull, 75, 424-429.

Hammer, R., Diesendruck, G., Weinshall, D., Hochstein, S. (2009). The development of category learning strategies: what makes the difference? Cognition, 112 (1), 105–119.

Herzog, M. H., & Fahle, M. (2002). Effect of Grouping in contextual modulation. Nature, 415(6870), 433-436.

Kurtzi, Z. (2010). Visual learning for perceptual and categorical decisions in the human brain. Vision Res 50(4), 433-440.

Nosofsky, R. M., & Palmeri, T. J. (1996). Learning to classify integral-dimension stimuli. Psychonomic Bull & Rev, 3, 222-226.

Rehder, B., & Hoffman, A. B. (2005). Thirty-something categorization results explained: Selective attention, eyetracking, and models of category learning. JEP: LMC, 31(5), 811-829.

Schoups, A., Vogel, R., & Orban, G. (2001). Practicing orientation identification improves orientation coding in V1 neurons. Nature, 412(6846), 549-553.

Seitz, A. R., Kim, D., & Watanabe, T. (2009). Rewards evoke learning of unconsciously processed visual stimuli in adult humans. Neuron, 61(5), 700-707.

Acknowledgments